



TITLE:

部分観測可能なマルコフ過程におけるジョブ・サーチについて (不確実性と意思決定数理の諸問題)

AUTHOR(S):

中井, 達

CITATION:

中井, 達. 部分観測可能なマルコフ過程におけるジョブ・サーチについて (不確実性と意思決定数理の諸問題). 数理解析研究所講究録 2004, 1373: 17-25

ISSUE DATE:

2004-05

URL:

<http://hdl.handle.net/2433/25521>

RIGHT:

部分観測可能なマルコフ過程におけるジョブ・サーチについて

中井 達 (Tōru Nakai)

九州大学大学院経済学研究院
(Faculty of Economics, Kyushu University)

1 はじめに

不完備情報の多段決定問題を、部分観測可能なマルコフ過程における決定問題としてモデル化し、とくにジョブ・サーチへの応用を考える。ジョブ・サーチにおいては、Lippman and MacCall [5] において、動的な経済 (dynamic economy) におけるものを考えているが、そこでは状態についての完全情報モデルであり、決定時点での状態を知ることができた。ここでは、不完備情報モデルとして部分観測可能なマルコフ過程におけるジョブ・サーチについて、最適政策とマルコフ過程の状態の推移に関連する性質をみる。このジョブ・サーチは最適停止問題の一つであり、Lippman and MacCall [5] のモデルのように、経済の状態をいくつかのクラスに分けて考えるが、この経済の状態がこれらのクラスのどこに属しているかは、一般的には未知であることが多く、このようなジョブ・サーチをモデル化する。このような、状態に関する情報は、何らかの形で予め知ることができるものとする。また、ジョブ・サーチにおいては、仕事から得られる賃金などは、これらの経済の状態に依存すると考えられるから、経済の状態に関する情報は、このような賃金などから得られると考えて良い。このようなジョブ・サーチにおいては、状態を知ることができる場合には、その最適政策は取りうる賃金の集合を reservation wage (留保賃金) と呼ばれる値によって分割される 2 つの互いに素な集合によって決まることが知られている。しかし、不完備情報のジョブ・サーチでは、この性質が必ずしも成り立たない。Nakai [12] において、部分観測可能なマルコフ過程における、これらの性質が調べられている。

観測できない状態に関する情報は、状態空間上の確率分布で表され、学習過程はベイズの定理にしたがうものとする。ここでは、 TP_2 (total positive of order two) として知られている性質を用いて、事前情報、事後情報、最適政策やマルコフ過程の状態の推移に関連する性質の関係をみる。この TP_2 については、Karlin and McGregor [3]、Karlin [2]、Karlin and Rinott [4] などでも、確率過程との関連について研究されている。

まずはじめに、動的な経済 (dynamic economy) におけるジョブ・サーチを考え、2 節で状態を直接知ることができる場合の結果を簡単にまとめる。3 節では、状態が部分観測可能なマルコフ過程にしたがって状態が推移するときに、状態の推移確率に関する性質を考え、合わせて最適政策にしたがった場合についても考える。また、ベイズの定理にしたがった学習過程を用いる場合についても、状態の推移確率の性質をみる。ところで、状態空間を $[0, S]$ とするマルコフ過程で、状態 S が債務不履行の状態を表すとすれば、これらの確率の性質をみることは、破産する確率の性質をみることになる。

2 動的な経済 (dynamic economy) におけるジョブ・サーチ

2.1 最適政策と期待利得

状態空間を $[0, S]$ とし、任意の状態 $s \in [0, S]$ に対して、状態空間 $[0, S]$ 上の確率分布の確率密度関数を $p_s = (p_s(t))_{t \in [0, S]}$ とする。ここで、 $p_s(t) \geq 0$ であり $\int_0^S p_s(t) dt = 1$ である ($t \in [0, S]$)。このとき、これらの確率密度関数を $p_s = (p_s(t))_{t \in [0, S]}$ は、状態が $s \in [0, S]$ のときのマルコフ過程の推移法則を表し、 $P = (p_s(t))_{s, t \in [0, S]}$ を推移確率密度とする。マルコフ過程の状態 s が経済の状態を表すとき、この状態に依存する賃金を表す確率変数を X_s とする ($s \in [0, S]$)。ジョブ・サーチとは、期待賃金を最大にする最適政策を求めることである。いま、ある人が仕事を探していて、費用 c を支払って一つの仕事が紹介され、合わせて m 個の仕事が出現するまで続ける。ここでは、採択しなかった仕事は再び現れることはない、リコールがない場合を考える。現れる仕事の賃金は、マルコフ過程の状態に依存する。状態空間が $\{1, 2, \dots, n\}$ のマルコフ連鎖の場合には、Lippman and MacCall [5] で、

(1) X_i は i に関して確率的に減少する。すなわち、任意の x に対して $F_1(x) \geq F_2(x) \geq \dots \geq F_n(x)$ であり、(2) $\sum_{j=k}^n p_{ij}$ は、任意の $k(k=1, 2, \dots, K)$ に対して i に関して増加するときに議論している。

ここでは、状態を直接知ることができない場合を扱うので、推移確率密度と確率変数 X_s ($s \in [0, S]$) の分布に関して、仮定 1 と 2 にもとで考える。この点が、上記の場合と異なってくる。状態空間が $[0, S]$ で、推移確率密度が $P = (p_s(t))_{s,t \in [0, S]}$ のとき、議論を簡単にするために確率変数 X_s は絶対連続で、密度関数 $f_s(x)$ を持つとする ($s \in [0, S]$) が、Nakai [10] で考えたようにこの仮定は一般化でき、いろいろな応用が考えられる (Nakai [7, 8, 9] など)。また、定義 1 において、全順序 \geq が定義された完備で可分な距離空間上の確率変数のあいだに、尤度比を用いて確率的な順序関係を導入する。

定義 1 確率変数 X と Y が、それぞれ密度関数 $f(x)$ と $g(x)$ を持ち、 $x \geq y$ となる任意の x と y に対して $f(y)g(x) \leq f(x)g(y)$ であれば、 X は Y より尤度比の意味で大きいといい、 $X \succeq Y$ と表す。

定義 2 関数 $P = (p_s(t))_{s,t \in [0, S]}$ が、 $s \leq t$ および $u \leq v$ となる任意の s, t, u と v に対して ($s, t, u, v \in [0, S]$)、 $p_s(u)p_t(v) \geq p_t(u)p_s(v)$ 、すなわち $\begin{vmatrix} p_s(u) & p_s(v) \\ p_t(u) & p_t(v) \end{vmatrix} \geq 0$ のとき、この P を TP_2 という。

定義 1 による順序が半順序となっていることは簡単に示され、この順序もまた TP_2 とよぶ。ここでは、確率変数 $\{X_s\}_{s \in [0, S]}$ は非負の実数値を取るものとし、2 つの仮定 (仮定 1 と 2) を設ける。これらの仮定は、学習過程としてベイズ学習を用いるために必要となるものである。

仮定 1 確率変数 $\{X_s \mid s \in [0, S]\}$ に対して、 $s \leq t$ ならば $X_s \succeq X_t$ である ($s, t \in [0, S]$)。すなわち、 X_s は s に関して尤度比の意味で減少する。

仮定 2 推移確率密度 $P = (p_s(t))_{s,t \in [0, S]}$ は TP_2 である。

仮定 1 において、 $s \leq t$ となる s と t に対して $X_s \succeq X_t$ だから ($s, t \in [0, S]$)、 $x > y$ ならば $f_s(y)f_t(x) \leq f_s(x)f_t(y)$ であるから、 s の値が大きくなるにつれ確率変数 X_s は小さい値を取りやすくなる。すなわち、状態 0 が最も良く、 \dots 、状態 S が最も悪いクラスである。仮定 2 は、このマルコフ過程に対して TP_2 として知られ、現在の状態から良いクラスの状態へ推移する確率は、現在の状態が良くなるにつれて大きくなる。この仮定から、状態を表す s が大きくなるにつれ、悪い状態へ推移する確率が増加する。

いま、 n 個の仕事が残っていて、直面している仕事からの賃金が x のとき、この仕事を採択すれば利得 $u_n(x)$ が得られる。このとき、状態が s で、直面している仕事からの賃金が x のとき、このジョブ・サーチの状態を (s, x) という。また、次の仕事を探すためには費用 c が必要であり、割引率を $0 < \beta < 1$ とする。ここで、 $v_n(s, x)$ を、 n 個の仕事が残っていて、ジョブ・サーチの状態が (s, x) のとき、最適に振る舞って得られる β で割り引いた期待利得とすれば、最適性の原理 (Ross [13] など) から、 $v_n(s, x)$ はつぎの最適方程式を満足する。

$$v_n(s, x) = \max \left\{ u_n(x), -c + \beta \int_0^S p_s(t) dt \int_0^\infty v_{n-1}(t, y) dF_t(y) \right\} \quad (1)$$

ここで、 $v_1(s, x) = u_1(x)$ とする。また、 $u_n(x)$ は x と n に関する増加関数とする。たとえば、 $u_n(x) = \frac{1 - \delta^n}{1 - \delta} x$ は条件を満足するが、これは資産 x を年利率 γ で n 年間預金したときの元利合計となっている ($\delta = 1 + \gamma$)。状態を知ることができるときには、最適政策は可能な仕事の賃金の集合を 2 つの互いに素な集合に分割することによって定まり、 n 個の仕事が残っていて、ジョブ・サーチの状態が (s, x) のとき、これらの集合は reservation wage $\alpha_n(s)$ により定まる。帰納法を用いれば、この $\alpha_n(i)$ が補題 1 を満足し、 β で割り引いた期待利得 $v_n(i, x)$ は補題 2 を満たす。

補題 1 任意の状態 $s \in [0, S]$ と正整数 n に対して、 $\alpha_{n+1}(s) \geq \alpha_n(s)$ であり、 $s < t$ となる任意の状態 $s, t \in [0, S]$ と正整数 n に対して、 $\alpha_n(s) \geq \alpha_n(t)$ である。

補題 2 任意の状態 $s \in [0, S]$ と正整数 n に対して、 $v_{n+1}(s, x) \geq v_n(s, x)$ かつ $v_{n+1}(s, x) \geq v_{n+1}(t, x)$ である ($x > 0, s < t$)。また、 $x > y$ ならば、 $v_{n+1}(s, x) \geq v_{n+1}(s, y)$ である。

これら2つの補題は、仮定1と2および $u_n(x)$ に関する仮定の下で、 n に関する帰納法を用いて示すことができる。いっぽう、 $u \leq v$ となる u と v に対して($u, v \in [0, S]$)、 P が仮定2を満たすことから $\begin{vmatrix} p_s(u) & p_s(v) \\ p_t(u) & p_t(v) \end{vmatrix} \geq 0$ となる。ここで、 $p_s = (p_s(u))$ および $p_t = (p_t(u))$ とおけば、任意の $s, t \in [0, S]$ に対して $p_t \succeq p_s$ であり、つぎの性質が成り立つ。

補題 3 $h(x)$ が x の非減少の非負関数とする。 $\mu \succeq \nu$ のとき、仮定1と2のもとで、

$$\int_0^\infty h(x) dF_\mu(x) \leq \int_0^\infty h(x) dF_\nu(x)$$

となる ($\mu, \nu \in S$)。

補題 4 $h(i, x)$ が i の減少する非負関数であり、 x の増加関数のとき、 $\mu \succeq \nu$ まらば ($\mu, \nu \in S$)、仮定1と2のもとで

$$\int_0^\infty \int_0^S h(s, x) \nu(s) f_s(x) dx ds \geq \int_0^\infty \int_0^S h(s, x) \mu(s) f_s(x) dx ds$$

である。

一般的に、つぎの性質が Nakai [6] と同じようにして求められる。すなわち、

$\int_0^S p_s(u) du \int_0^\infty v_n(u, y) dF_u(y) = \int_0^\infty \int_0^S p_s(u) v_n(u, y) f_u(y) dy du$ だから、 $s < t$ のとき $p_t \succeq p_s$ であることと補題4から

$$\int_0^S p_s(u) du \int_0^\infty v_n(u, y) dF_u(y) \geq \int_0^S p_t(u) du \int_0^\infty v_n(u, y) dF_u(y)$$

となる。

2.2 状態への推移確率

つぎに、マルコフ過程の状態を知ることができるとき、 n 期間後に状態が t となる確率を考える。はじめに、状態の推移のみに着目し、マルコフ過程の状態が s のとき、 $\bar{p}_{s,n}(t)$ を n 期間後における状態を表す確率変数の確率密度とする ($s, t \in [0, S], n = 1, 2, \dots$)。このとき、この密度関数 $\bar{p}_{s,n}(t)$ が、初期条件を $\bar{p}_{s,1}(t) = p_s(t)$ とする再帰式 $\bar{p}_{s,n}(t) = \int_0^S p_s(u) \bar{p}_{u,n-1}(t) du$ を満たすことは簡単にわかる。いま、関数 $\bar{P}_n = (\bar{p}_{s,n}(t))_{s,t \in [0,S]}$ に対して、 $\bar{P}_1 = P$ であり $\bar{P}_n = \langle P, \bar{P}_{n-1} \rangle$ となっている。ここで、2つの関数 $P = (p_s(t))_{s,t \in [0,S]}$ と $Q = (q_s(t))_{s,t \in [0,S]}$ に対して、

$$\langle P, Q \rangle = \left(\int_0^S p_s(u) q_u(t) du \right)_{s,t \in [0,S]}$$

と定義する。このとき、つぎの性質が成り立つ。

補題 5 $P = (p_s(t))_{s,t \in [0,S]}$ と $Q = (q_s(t))_{s,t \in [0,S]}$ が TP_2 であれば、 $\langle P, Q \rangle$ もまた TP_2 である。

このとき、仮定2から $P = (p_s(t))_{s,t \in [0,S]}$ が TP_2 だから、 $\bar{P}_{n-1} = (p_{s,n-1}(t))_{s,t \in [0,S]}$ が TP_2 であることが示されれば、 n に関する帰納法と補題5より、 $\bar{P}_n = \langle P, \bar{P}_{n-1} \rangle = (\bar{p}_{s,n}(t))_{s,t \in [0,S]}$ が TP_2 となる。

ここで、確率密度 $\bar{p}_{s,n}(t)$ は、マルコフ過程の状態が s のとき、部分観測可能なマルコフ過程にしたがって状態が推移して、 n 期後の状態を表す確率変数の確率密度であった。つぎに、このマルコフ過程にしたがって状態が推移するジョブ・サーチにおいて、最適政策にしたがったときに、同様の確率を考える。現在の状態が s で、 n 個の仕事が残っているとき、 $\bar{p}_{s,n,m}(t)$ を最適政策にしたがったとき、 m 期後の状態を表す確率変数の確率密度とする ($s, t \in [0, S]$ および $m \leq n, n, m = 1, 2, \dots$)。状態を直接に知ることができるジョブ・サーチにおいては、最適政策は reservation wages $\alpha(s, n)$ によって定まり、 $F_s(\alpha(s, n))$ は、状態が s で、 n 個の仕事が残っているときに、直面している仕事を採択し

ない確率である。したがって、 $\bar{p}_{s,n,m} = (\bar{p}_{s,n,m}(t))_{t \in [0,S]}$ が、初期条件を $\bar{p}_{s,n,1} = (\bar{p}_{s,n,1}(t))_{t \in [0,S]}$ とする再帰式

$$\bar{p}_{s,n,m}(t) = F_s(\alpha(s,n)) \int_0^S p_s(x) \bar{p}_{x,n-1,m-1}(t) dx \quad (2)$$

を満たす。ただし、 $\bar{p}_{s,n,1}(t) = F_s(\alpha(s,n))p_s(t)$ とする。ここで、

$$\bar{P}_{n,m} = (\bar{p}_{s,n,m})_{s \in [0,S]} = (\bar{p}_{s,n,m}(t))_{s,t \in [0,S]}$$

とおけば、任意の $n(>0)$ に対して $\bar{P}_{n,1} = (F_s(\alpha(s,n))p_s)_{s \in [0,S]} = (F_s(\alpha(s,n))p_s(t))_{s,t \in [0,S]}$ であることと (2) 式より

$$\bar{P}_{n,m} = (F_s(\alpha(s,n))\langle P, \bar{P}_{n-1,m-1} \rangle_s)_{s \in [0,S]} \quad (3)$$

となる。このとき、補題 5 からつぎの性質が示される。

系 1 $P = (P_s)_{s \in [0,S]}$ が TP_2 であり、 $d(s)$ を s の関数とする。このとき、 $Q = (d(s)P_s)_{s \in [0,S]} = (d(s)p_s(t))_{s,t \in [0,S]}$ もまた TP_2 である

これらの $\bar{P}(n,m)$ は、つぎの性質を満足する。

命題 1 $\bar{P}_{n,m} = (\bar{p}_{s,m,n})_{s \in [0,S]} = (\bar{p}_{s,m,n}(t))_{s,t \in [0,S]}$ は TP_2 である。

証明: m に関する帰納法を用いる。 $m = 1$ のとき、系 1 より $\bar{P}_{n,1} = (F_s(\alpha(s,n))p_s)_{s \in [0,S]} = (F_s(\alpha(s,n))p_s(t))_{s,t \in [0,S]}$ は TP_2 である。つぎに、 $\bar{P}_{n,m}$ が m 以下の値に対して、 TP_2 とする。 $P = (p_s(t))_{s,t \in [0,S]}$ と $\bar{P}_{n-1,m-1}$ が TP_2 だから、系 1 より $\langle P, \bar{P}_{n-1,m-1} \rangle$ もまた TP_2 である。したがって、

$$\bar{P}_{n,m} = (F_s(\alpha(s,n))\langle P, \bar{P}_{n-1,m-1} \rangle_s)_{s \in [0,S]}$$

もまた TP_2 である。□

3 不完備情報のジョブ・サーチ

3.1 最適政策と期待利得

前節で考えたマルコフ過程におけるジョブ・サーチで、状態を直接知ることができない場合、すなわち部分観測可能なマルコフ過程でのジョブ・サーチを考える。状態に関する情報は、状態空間 $[0, S]$ 上の確率分布 μ で表され、 S を情報全体の集合とすれば、

$$S = \left\{ \mu = (\mu(s))_{s \in [0,S]} \mid \int_0^S \mu(s) = 1, \mu(s) \geq 0 (s \in [0, S]) \right\}$$

となる。これらの確率分布 μ は密度関数を持つものとする。また、 S に含まれる情報のあいだに、定義 1 によって尤度比を用いた順序関係を導入する。すなわち、状態空間 $[0, S]$ 上の 2 つの確率分布 μ, ν に対して、 $\mu(t)\nu(s) \geq \mu(s)\nu(t)$ が任意の $s, t (s \leq t, s, t \in [0, S])$ に対して成り立ち、少なくとも 1 つの s と t の組み合わせに対して $\mu(t)\nu(s) > \mu(s)\nu(t)$ のとき、 μ は ν より尤度比の意味で大きいといい、簡単に $\mu \succ \nu$ と表す。この順序関係は半順序であり、簡単に TP_2 という。定義 1 より、 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、 t が大きくなるにしたがって、状態 t における密度の比 $\frac{\mu(t)}{\nu(t)}$ は $\nu(t) \neq 0$ の範囲で増加する。この関係は、一般的な部分観測可能なマルコフ過程においても定義でき、詳しくは多段決定問題への応用を含めて Nakai [10, 11] にある。

観測できない状態に関して、その状態に関する情報を得るための情報過程が存在すると考える。ここでは、確率変数 $\{X_s \mid s \in [0, S]\}$ が未知の状態に依存する仕事の賃金を表すから、これらの確率変数を観測することが情報過程にあたる。したがって、観測できない状態に関して、この賃金を用いて情報を改良する。このとき、事前情報が μ のとき、賃金を知って、学習過程としてベイズの定理を用いることによって、情報を改良する。そのあとで、マルコフ過程の状態が推移し、新しい状態へ移ると考える。この順番は、逆にしても可能であるが、基本的には同様である。いま、直面する仕事の賃金が x のとき、状態についての新しい情報をベイズの定理を用いて $\mu(x) = (\mu(x, s))_{s \in [0,S]}$ と改良し、そのあと推移確率密度 P にしたがって状態が推移する。このとき、つぎの決定時点における事前情報を $\mu(x) = (\mu(x, s))_{s \in [0,S]}$ とする。つぎに、 x と s の関数 $h(x, s)$ に対して、定義 3 より単調性を定義する。

定義 3 $s \in [0, S]$ と $x \in \mathbb{R}$ の非負関数 $h(x) = (h(x, s))_{s \in [0, S]}$ に対して、 $x < y$ ならば $s \leq t$ となる任意の t と s について $(s, t \in [0, S])$ 、 $h(x) \succeq h(y)$ (または $h(y) \succeq h(x)$) すなわち $h(x, t)h(y, s) \geq h(x, s)h(y, t)$ (または $h(x, t)h(y, s) \leq h(x, s)h(y, t)$) とする。このとき、関数 $h(x, s)$ を x の減少関数 (または増加関数) という。

確率変数 $\{X_s \mid s \in [0, S]\}$ は密度関数 $\{f_s(x) \mid s \in [0, S]\}$ を持ち、仮定 1 を満たすから、 $f(x) = (f_s(x))_{s \in [0, S]}$ とおけば $f(y) \succeq f(x)$ となる。すなわち、 $x > y$ ならば、 $s \leq t$ となる任意の s と t に対して $(s, t \in [0, S])$ 、 $f_s(y)f_t(x) \leq f_s(x)f_t(y)$ である。したがって、 $f(x)$ は x に関する増加関数である。

事前情報 μ と事後情報 $\overline{\mu(x)}$ については、仮定 1 と 2 のもとで、次のような補題 6 が成り立ち、Nakai [10] などと求められている。

補題 6 $\mu \succ \nu$ ならば、任意の x に対して $\mu(x) \succ \nu(x)$ かつ $\overline{\mu(x)} \succ \overline{\nu(x)}$ である。任意の μ に対して、 $\mu(x)$ と $\overline{\mu(x)}$ は x の減少関数である。

補題 6 から、事前情報 μ のあいだの順序関係 (定義 1) は改良された情報 $\mu(x)$ と事後情報 $\overline{\mu(x)}$ のあいだでも保存される。さらに、同じ事前情報 μ であっても、仕事から得られる賃金 x が大きくなれば、事後情報 $\overline{\mu(x)}$ は尤度比の意味で悪くなる (定義 1)。

マルコフ過程の未知の状態に関する事前情報が μ であるジョブ・サーチを考える。いま、 n 個の仕事が残っていて、直面している仕事の賃金が x のとき、 $v_n(\mu, x)$ を、最適政策を用いたときの β で割引された総期待利得とする ($0 < \beta < 1$)。最適性の原理より、この $v_n(\mu, x)$ は次の再帰方程式を満足する。

$$v_n(\mu, x) = \max \left\{ u_n(x), c + \beta \int_0^\infty v_{n-1}(\overline{\mu(x)}, y) dF_{\overline{\mu(x)}}(y) \right\} \quad (4)$$

ただし、 $F_\mu(x) = \int_0^x \mu(s)F_s(x)$ は weighted distribution function (De Vylder[1]) であり、 $v_1(\mu, x) = E_\mu[u_1(X)] = \int_0^\infty u_1(x)dF_\mu(x)$ とする。また、状態空間が $S = [0, S]$ のとき、推移確率密度 P に対して $p_S(t) = I_S(t)$ であり、確率 1 で $X = 0$ とすれば、状態 S は債務不履行の状態を表すと考えられる。ここで、 $I_S(t)$ は t の indicator function である。

ここで

$$S(\mu, n) = \left\{ x \mid u_n(x) \geq c + \beta \int_0^\infty v_{n-1}(\overline{\mu(x)}, y) dF_{\overline{\mu(x)}}(y) \right\}$$

および $C(\mu, n) = S(\mu, n)^c$ とおけば、 $S(\mu, n)$ と $C(\mu, n)$ は、それぞれジョブ・サーチにおける停止領域と継続領域を表す。ここで、 $u_n(x)$ は x の増加関数であり、 $\overline{\mu(x)}$ は x の減少関数であることに注意する。すなわち、 $x > y$ ならば $\overline{\mu(y)} \succeq \overline{\mu(x)}$ である。また、被積分関数 $v_{n-1}(\overline{\mu(x)}, z)$ が z の増加関数であり、 μ の減少関数のとき、補題 3 より $x > y$ ならば

$$\int_0^\infty v_{n-1}(\overline{\mu(x)}, z) dF_{\overline{\mu(x)}}(z) \geq \int_0^\infty v_{n-1}(\overline{\mu(y)}, z) dF_{\overline{\mu(y)}}(z)$$

となる。したがって、これら 2 つの領域 $S(\mu, n)$ と $C(\mu, n)$ に関して、(4) 式から次の性質が得られる。

補題 7 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、 $S(\nu, n) \subset S(\mu, n)$ および $S(\mu, n+1) \subset S(\mu, n)$ である。

任意の μ と $n \geq 1$ に対して、 $S(\mu, n) \cup C(\mu, n) = \mathbb{R}_+$ であり、 $S(\mu, n) \cap C(\mu, n) = \emptyset$ であるから、この補題から $C(\mu, n) \subset C(\nu, n)$ および $C(\mu, n) \subset C(\mu, n+1)$ となる。また、 $v_n(\mu, x)$ は次の性質を持つ。

補題 8 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、 $v_n(\mu, x) \leq v_n(\nu, x)$ である。また、 $x > y$ ならば、 $v_{n+1}(\mu, x) \geq v_n(\mu, x)$ および $v_n(\mu, x) \geq v_n(\mu, y)$ である。

これらの性質は Nakai [10] などと同じように、 n に関する帰納法で示される。

3.2 状態への推移確率—不完備情報の場合

2.2 節と同じように、状態が部分観測可能なマルコフ過程に従って推移する場合、仮定 1 と 2 のもとで、 n 期間後に状態が t となる確率を考える。はじめに、順序立てて考えるために、これらの確率を決定と未知の状態に関する学習過程を除いて考える。未知の状態に関する事前情報が μ のとき、 $\bar{P}_m(\mu)$ を m 期間後の状態を表す確率変数の確率密度とする。初期条件として、 $m = 1$ のときは $\bar{P}_{\mu,1} = (\bar{P}_{\mu,1}(t))_{t \in [0,S]}$ であり、 $\bar{P}_1(\mu)_t = \int_0^S \mu(s)p_s(t)ds = \langle \mu, P \rangle(t)$ となる。前節と同じように、 $\mu = (\mu(s))_{s \in [0,S]}$ と $P = (p_s(t))_{s,t \in [0,S]}$ に対して、 $\langle \mu, P \rangle$ を $\langle \mu, P \rangle = (\langle \mu, P \rangle(t))_{t \in [0,S]}$ および

$$\langle \mu, P \rangle(t) = \int_0^S \mu(s)p_s(t)ds$$

とする。このとき、 $\langle \langle \mu, P \rangle, Q \rangle = \langle \mu, \langle P, Q \rangle \rangle$ であることは明らかである。さらに、 $P = (p_s(t))_{s,t \in [0,S]}$ に対して、 P^n を $P^1 = P$ および $P^n = \langle P, P^{n-1} \rangle$ で定義する。このとき、 $\bar{\mu} = \langle \mu, P \rangle$ および $\bar{\mu}(x) = \langle \mu(x), P \rangle$ となる。この関係式から $m = 2$ に対して $\bar{P}_{\mu,2} = \bar{P}_{\bar{\mu},1} = \langle \bar{\mu}, P \rangle = \langle \mu, P^2 \rangle$ であり、 $\bar{P}(\mu, m)$ の再帰関係式はつぎのようになる。

$$\bar{P}_{\mu,m} = \bar{P}_{\bar{\mu},m-1} = \bar{P}_{\langle \mu, P \rangle, m-1} = \langle \langle \mu, P \rangle, P^{m-1} \rangle = \langle \mu, P^m \rangle \quad (5)$$

ここで、 P が TP_2 だから、 m に関する帰納法より $\bar{P}_m(\mu) = \langle \mu, P^m \rangle$ もまた TP_2 であることがわかる。よって、つぎの性質が成り立つ。

補題 9 μ と ν であり $(\mu, \nu \in S)$ 、 P が TP_2 ならば $(\mu, \nu \in S)$ 、 $\langle \mu, P \rangle \succeq \langle \nu, P \rangle$ である。

(5) 式より、補題 9 からつぎの性質が得られる。

命題 2 μ と ν ならば $(\mu, \nu \in S)$ 、 $\bar{P}_{\mu,m} \succeq \bar{P}_{\nu,m}$ である。

つぎに、同様の確率を、決定を除いて考える。すなわち、直面している仕事の賃金 x を用いて、未知の状態に関する学習過程を考慮する。事前情報が μ のとき、はじめに状態に依存する x の値を知って、ベイズの定理に従って情報を改良する。そのあと、つぎの期へ進み、推移確率密度 $P = (p_s(t))_{s,t \in [0,S]}$ にしたがって状態の推移が起こる。したがって、状態に関する事前情報が μ であるというとき、推移は終わっているものとする。状態に関する事前情報が μ のとき、 $\hat{P}_{\mu,m}(t)$ を m 期間後の状態を表す確率変数の確率密度とし $(t \in [0, S])$ 、 $\hat{P}_{\mu,m} = (\hat{P}_{\mu,m}(t))_{t \in [0,S]}$ とする。

簡単のために、関数 $u(x) = (u_t(x))_{t \in [0,S]}$ について、任意の s に対して $\int_a^b u_s(x)dF(x)$ が存在すれば $(s \in [0, S])$ 、 $\int_a^b u(x)dF(x)$ を

$$\int_a^b u(x)dF(x) = \left(\int_a^b u_t(x)dF(x) \right)_{t \in [0,S]}$$

と表す。

未知の状態に関する事前情報が μ のとき、 $\hat{P}_{\mu,1} = (\hat{P}_{\mu,1}(t))_{t \in [0,S]}$ がつぎの時点での状態を表す確率変数の確率密度だから、

$$\hat{P}_{\mu,1} = \int_0^\infty \langle \mu(x), P \rangle dF_\mu(x) = \int_0^\infty \bar{\mu}(x) dF_\mu(x)$$

である。ある時点での事前情報が μ で、直面する仕事の賃金が x のとき、つぎの期での事前情報が $\bar{\mu}(x)$ であった。 $\hat{P}_{\mu,2}$ が 2 期間後の状態を表す確率変数の確率密度だから、 $m = 1$ のときと同様に、 $\hat{P}_{\mu,2} = \int_0^\infty \hat{P}_{\bar{\mu}(x),1} dF_\mu(x)$ となる。同じように事前情報が μ のとき、 $\hat{P}_{\mu,m}$ が m 期間後の状態を表す確率変数の確率密度だから、 $\hat{P}_{\mu,m}$ は (6) 式を満足する。

$$\hat{P}_{\mu,m} = \int_0^\infty \hat{P}_{\bar{\mu}(x),m-1} dF_\mu(x), \quad (6)$$

ここで、 $\hat{P}_{\mu,1} = \int_0^\infty \overline{\mu(x)} dF_\mu(x)$ とする。これらの確率密度の性質をみるため、定義 4 により順序関係を入れる。

定義 4 x の非負関数 $g(x) = (g(x,s))_{s \in [0,S]}$ と $h(x) = (h(x,s))_{s \in [0,S]}$ が、 $s \leq t$ となる任意の s と t に対して $(s,t \in [0,S])$ 、 $g(x,t)h(x,s) \geq g(x,s)h(x,t)$ ならば、 $g(x)$ は $h(x)$ より TP_2 の意味で大きいといい、 $g(x) \succeq h(x)$ と表す。

このとき、つぎの 2 つの性質が成り立つ。

補題 10 非負関数 $g(x) = (g(x,s))_{s \in [0,S]}$ と $h(x) = (h(x,s))_{s \in [0,S]}$ が x の減少関数であり、 $g(x) \succeq h(x)$ ならば、

$$\int_0^\infty g(x) dF(x) \succeq \int_0^\infty h(x) dF(x)$$

である。

補題 11 $\mu \succeq \nu$ であり $(\mu, \nu \in S)$ 、集合値関数 $h(x)$ が x の減少関数であれば、仮定 1 と 2 のもとで

$$\int_0^\infty h(x) dF_\mu(x) \succeq \int_0^\infty h(x) dF_\nu(x)$$

である。

つぎの 2 つの性質は補題 10 と 11 より導かれる。

系 2 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、仮定 1 と 2 のもとで

$$\int_0^\infty \overline{\mu(x)} dF(x) \succeq \int_0^\infty \overline{\nu(x)} dF(x)$$

となる。

系 3 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、 μ の増加かつ x の非増加関数 $h(\mu, x)$ に対して、仮定 1 と 2 のもとで

$$\int_0^\infty h(\mu, x) dF_\mu(x) \succeq \int_0^\infty h(\nu, x) dF_\nu(x)$$

となる。

$\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、補題 6 より $\overline{\mu} \succeq \overline{\nu}$ および $\overline{\mu(x)} \succeq \overline{\nu(x)}$ となる。したがって、 $\hat{P}_{\mu,m}$ はつぎの性質を持つ。

命題 3 $\mu \succeq \nu$ ならば $(\mu, \nu \in S)$ 、 $\hat{P}_{\mu,m}$ は μ の増加関数である。すなわち、 $\hat{P}_{\mu,m} \succeq \hat{P}_{\nu,m}$ である。

証明: m に関する帰納法を用いる。 $m=1$ のとき、 $\hat{P}_{\mu,1} = \int_0^\infty \overline{\mu(x)} dF_\mu(x)$ であり、 $\mu \succeq \nu$ ならば、系 2 から $\hat{P}_{\mu,1} \succeq \hat{P}_{\nu,1}$ となる。また、 $\overline{\mu(x)} \succeq \overline{\nu(x)}$ だから、 $\hat{P}_{\overline{\mu(x)},1} \succeq \hat{P}_{\overline{\nu(x)},1}$ である。系 3 より

$$\hat{P}_{\mu,2} = \int_0^\infty \hat{P}_{\overline{\mu(x)},1} dF_\mu(x) \succeq \int_0^\infty \hat{P}_{\overline{\nu(x)},1} dF_\mu(x) = \hat{P}_{\nu,2}$$

となり、 $\hat{P}_{\mu,2} \succeq \hat{P}_{\nu,2}$ が示される。

帰納法の仮定から、 $\mu \succeq \nu$ ならば、 $\hat{P}_{\mu,m-1} \succeq \hat{P}_{\nu,m-1}$ となる。 $\overline{\mu(x)} \succeq \overline{\nu(x)}$ より $\hat{P}_{\overline{\mu(x)},m-1} \succeq \hat{P}_{\overline{\nu(x)},m-1}$ である。したがって、系 3 から

$$\begin{aligned} \hat{P}_{\mu,m} &= \int_0^\infty \hat{P}_{\overline{\mu(x)},m-1} dF_\mu(x) \\ &\succeq \int_0^\infty \hat{P}_{\overline{\nu(x)},m-1} dF_\mu(x) \\ &\succeq \int_0^\infty \hat{P}_{\overline{\nu(x)},m-1} dF_\nu(x) = \hat{P}_{\nu,m}, \end{aligned}$$

となり、この性質が示される。□

最後に、同様の確率を、決定と学習過程を含めて考える。すなわち、事前情報が μ のとき、直面している仕事の賃金 x を知って、未知の状態について学習を行い、この情報のもとで、この仕事に就くかどうかを決定する。もし、見送れば 1 期進み、マルコフ過程は推移確率密度 P にしたがって新しい状態へと推移し、状態についての事後情報は $\bar{\mu}(x)$ となる。

いま、 n 個の仕事が残っていて、状態に関する事前情報が μ のとき、直面している仕事の賃金を x とする。 $(\tilde{P}_{\mu,n,m}(t))_{t \in [0,S]}$ を最適政策にしたがったときの m 期間後の状態を表す確率変数の確率密度とする ($t \in [0, S], n, m = 1, 2, \dots, m \leq n$)。

これらの $(\tilde{P}_{\mu,n,m}(t))_{t \in [0,S]}$ を考えるために、はじめに $m = 1$ の場合を考える。直面している仕事の賃金 x を観測し、この仕事を採択しなかったとしよう。このとき、状態について改良した情報を簡単のために μ^* とおく。 $(\tilde{P}'_{\mu^*,n,1}(t))_{t \in [0,S]}$ を、推移をしたあとでの、つぎの期での状態を表す確率変数の確率密度とする ($s \in [0, S]$)。このとき、 $x \in C(\mu, n)$ のときにのみ、つぎの期へ進むから、 $\tilde{P}'_{\mu^*,n,1} = (\tilde{P}'_{\mu^*,n,1}(t))_{t \in [0,S]}$ は $\tilde{P}'_{\mu^*,n,1}(t) = \int_0^S \mu^*(s) p_s(t)$ と $\tilde{P}'_{\mu^*,n,1} = \langle \mu^*, P \rangle = \bar{\mu}^*$ を満たす。直面する仕事の賃金 x を観測したとき、改良した情報は $\mu(x)$ だから、

$$\tilde{P}_{\mu,n,1}(t) = \int_{C(\mu,n)} \tilde{P}'_{\mu(x),n,1}(t) dF_{\mu}(x) = \int_{C(\mu,n)} \int_0^S \mu(x)_s p_s(t) dF_{\mu}(x)$$

となる。したがって、 $\tilde{P}_{\mu,n,1} = (\tilde{P}_{\mu,n,1}(t))_{t \in [0,S]}$ と $\tilde{P}'_{\mu,n,1}$ には、

$$\tilde{P}_{\mu,n,1} = \int_{C(\mu,n)} \tilde{P}'_{\mu(x),n,1} dF_{\mu}(x) = \int_{C(\mu,n)} \bar{\mu}(x) dF_{\mu}(x)$$

の関係がある。

n 個の仕事が残っているとき、状態に関する事前情報が μ であれば、 $(\tilde{P}_{\mu,n,m}(t))_{t \in [0,S]}$ を、最適政策にしたがったときの m 期間後の状態を表す確率変数の確率密度とする ($t \in [0, S], n, m = 1, 2, \dots, m \leq n$)。このとき、新しい仕事が見れ、その賃金は状態に依存する。この値をもとに、この仕事に就くかどうかを決定する。ジョブ・サーチにおいては、 $x \in C(\mu, n)$ のとき、つぎの仕事へと進むから、 $\tilde{P}_{\mu,n,m} = (\tilde{P}_{\mu,n,m}(t))_{t \in [0,S]}$ が再帰方程式

$$\tilde{P}_{\mu,n,m}(t) = \int_{C(\mu,n)} \tilde{P}_{\mu(x),n-1,m-1}(t) dF_{\mu}(x) \quad (7)$$

を満たす。ここで、 $\int_{S(\mu,n)} dF_{\mu}(x)$ は直面する仕事に就く確率だから、 $\int_0^S \tilde{P}_{\mu,n,m}(t) dt \leq 1$ であることは明らかである。さらに、 $\mu \succeq \nu$ ならば $C(\mu, n) \subset C(\nu, n)$ となる。すなわち、見送ってつぎの仕事を探す確率は、 μ が増加するにしたがって減少する。いっぽう、より悪い状態へ推移する確率は、 μ が増加すれば増加する。マルコフ過程の状態を直接知ることができる場合には、性質 1 より、 $\bar{P}_{n,m} = (\bar{P}_{s,n,m})_{s \in [0,S]}$ は TP_2 であった。しかし、この確率 $(\tilde{P}_{\mu(x),n-1,m-1}(t))_{t \in [0,S]}$ が観測した x によって変化するので、この場合には $\tilde{P}_{\mu,n,m}$ がその様な性質を持つことを示すことは難しい。

なお、ここで用いた TP_2 の性質を用いて、いくつかの定義を行った (定義 1, 2, 3) が、ここで用いた補題や系は、これらの定義から仮定 1 と 2 のもとで導かれる。また、一般の場合には Nakai [10, 7, 6] などのように拡張できる。

References

- [1] De Vylder, F. (1983). Duality Theorem for Bounds in Integrals with Applications to Stop Loss Premiums, *Scandinavian Actuarial Journal*, 129–147.
- [2] S. Karlin, *Total Positivity*, Stanford University Press, Stanford, California (1968).
- [3] S. Karlin and J. L. McGregor, Classical Diffusion Process and Total Positivity, *Journal of Mathematical Analysis and Applications*, 1, 163–183 (1960).

- [4] S. Karlin and Y. Rinott, Total Positivity Properties of Absolute Value Multinomial Variables with Applications to Confidence Interval Estimates and Related Probabilistic Inequalities, *The Annals of Statistics*, **9**, 1035–1049 (1981).
- [5] S. A. Lippman and J. J. McCall, Job Search in a Dynamic Economy, *Journal of Economic Theory*, **12**, 365–390 (1976).
- [6] T. Nakai, The Problem of Optimal Stopping in a Partially Observable Markov process, *Journal of Optimization Theory and Applications*, **45**, 425–442 (1985).
- [7] T. Nakai, A Sequential Stochastic Assignment Problem in a Partially Observable Markov process, *Mathematics of Operations Research*, **11**, 230–240 (1986).
- [8] T. Nakai, An Optimal Selection Problem on a Partially Observable Markov process, *Stochastic Modelling in Innovative Manufacturing*, Lecture Notes in Economics and Mathematical Systems **445**, Eds. A. H. Christer, S. Osaki and L. C. Thomas, 140–154, Springer-Verlag, Berlin (1996).
- [9] T. Nakai, An Optimal Assignment Problem for Multiple Objects per Period – Case of a Partially Observable Markov process, *Bulletin of Informatics and Cybernetics*, **31**, 23–34 (1999).
- [10] T. Nakai, A Generalization of Multivariate Total Positivity of Order Two with an Application to Bayesian Learning Procedure, *Journal of Information & Optimization Sciences*, **23**, 163–176 (2002).
- [11] T. Nakai, A Generalized MTP_2 and a Sequential Stochastic Model on a Partially Observable Markov Process, *Probabilistic Methods in Discrete Mathematics—Proceedings of the Fifth International Petrozavodsk Conference*, (Eds. V.F. Kolchin, V.Ya. Kozlov, V.V. Mazalov, Yu.L. Pavlov and Yu.V. Prokhorov), VSP publishes, The Netherlands, 291–302 (2002).
- [12] T. Nakai, Properties of a Partially Observable Markov process for a Job Search Problem in a Dynamic Economy, *Proceedings of the 2nd Euro Japanese Workshop on Stochastic Modelling for Finance, Insurance, Production and Reliability*, (Eds. T. Dohi, N. Limonios and S. Osaki), Systems Reliability Engineering Laboratory, Hiroshima University, 340–349, 2002.
- [13] S. M. Ross, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, California (1970).